Markov chain Monte Carlo methods for the Box-Behnken designs and centrally symmetric configurations

Satoshi Aoki* , Takayuki Hibi†and Hidefumi Ohsugi‡

February 10, 2015

Abstract

We consider Markov chain Monte Carlo methods for calculating conditional p values of statistical models for count data arising in Box-Behnken designs. The statistical model we consider is a discrete version of the first-order model in the response surface methodology. For our models, the Markov basis, a key notion to construct a connected Markov chain on a given sample space, is characterized as generators of the toric ideals for the centrally symmetric configurations of root system D_n. We show the structure of the Gröbner bases for these cases. A numerical example for an imaginary data set is given.

1 Introduction

After the work by Diaconis and Sturmfels ([10]), a *Markov basis*, a key notion in the field of computational algebraic statistics, has attracted special attentions among researchers both in statistics and algebra. In this first work, they show the fundamental relation between the generators of toric ideals and the Markov bases and establish a procedure for sampling from discrete conditional distributions by constructing an irreducible Markov chain on a given sample space. By virtue of this relation, we can perform Markov chain Monte Carlo methods to estimate conditional p values for various statistical problems if we can obtain the generator of corresponding toric ideals. Readers can find various theoretical results on structure of Markov

*Graduate School of Science and Engineering (Science Course), Kagoshima University.

†Department of Pure and Applied Mathematics, Graduate School of Information Science and Technology, Osaka University

‡Department of Mathematical Sciences, School of Science and Technology, Kwansei Gakuin University

bases such as minimality or invariance, and Markov bases of important statistical models such as hierarchical models of multi-dimensional contingency tables in [2].

In parallel, it is also valuable to connect known classes of toric ideals to statistical models. Such a motivation yields attractive research topics from algebraic fields to statistics. For example, [5] shows the relation between the generator of the toric ideals for the Segre–Veronese configuration and the special independence models in the testing problems of group-wise selections. This result is further generalized to a class of configurations called nested configurations in [4]. As another example, relations between regular two-level fractional factorial designs and cut ideals are shown in [3]. The arguments in this paper comes from the same motivation to these works.

In this paper, we consider the statistical models corresponding to the algebraic object known as a *centrally symmetric configuration of root system* D_n. The notion of centrally symmetric configurations ([15]) is one of the new attractive topics in algebra since it yields many "toric rings" that have important algebraic properties (normal and Gorenstein). See [15]. On the other hand, Gröbner bases of the toric ideal arising from the configuration of D_n is studied in [14]. In addition, convex polytopes arising from the centrally symmetric configuration of D_n are studied in [8]. In this paper, we show that the centrally symmetric configuration corresponds to the first-order models for the symmetric designs of experiments for multi-level factors. As typical examples of such designs, we consider Box-Behnken designs in this paper. Markov chain Monte Carlo procedure in the framework of design of experiments is introduced in [7] and [6]. In these works, regular two-level and three-level designs are considered. However, non-regular designs are difficult to treat in general. In this paper, we present a new method for analyzing non-regular designs.

The construction of this paper is as follows. In Section 2, we review the Markov chain Monte Carlo methods for design of experiments. In Section 3, we give a definition of the centrally symmetric configuration and present statistical models. We also introduce the Box-Behnken designs and show that the model matrix of the first-order models for the Box-Behnken designs corresponds to the centrally symmetric configuration of root system D_n. In Section 4, we give the Gröbner bases of the centrally symmetric configuration of root system D_n. In Section 5, we give numerical example for an imaginary data set. Finally, we give some discussion in Section 6.

2 Markov chain Monte Carlo methods for design of experiments

In this section, we introduce Markov chain Monte Carlo methods for testing the fitting of the log-linear models for fractional factorial designs with count observations. We consider the designs with m controllable factors. For $j = 1, \dots, m$, write $A_j \in \mathbb{Q}$ as the level of the j-th factor. For example, if j-th factor has three levels,

it is common to write $A_j = \{-1, 0, 1\}$. The full factorial design $P \subset \mathbb{Q}^m$ is given as $P = A_1 \times \cdots \times A_m$ and the fractional factorial design F is a subset of P. Suppose there are k runs (i.e., points) in F. For convenience, we order the points of F appropriately and consider a $k \times m$ design matrix $D = (d_{ij})$, where d_{ij} is the level of j-th factor in i-th run for $i = 1, \dots, k,\ j = 1, \dots, m$.

We write the observations as $\mathbf{y} = (y_1, \dots, y_k)'$, where $'$ denotes the transpose. In this paper, we consider the Poisson sampling scheme, i.e., we suppose that the observations are counts of some events and the observation $\mathbf{y}$ is nonnegative integer vector. We also suppose that only one observation is obtained for each run. This is a natural setting because the set of the totals for each run is the sufficient statistics for the parameter in the Poisson sampling scheme. Therefore the observation $\mathbf{y}$ are realizations from k mutually independent Poisson random variables $Y_1, \dots, Y_k$ with the mean parameter $\lambda_i = E(Y_i),\ i = 1, \dots, k$.

We consider the log-linear model

$$\log \lambda_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_n x_{in},\ i = 1, \dots, k \tag{1}$$

for the parameter $\lambda_i, i = 1, \dots, k$, where x_{ij} is a j-th covariate for the i-th run and $n+1$ is the dimension of the parameter $\beta = (\beta_0, \beta_1, \dots, \beta_n)'$. If we write $x_{i0} = 1$ for $i = 1, \dots, k$, the log-linear model (1) is written as

$$(\log \lambda_1, \dots, \log \lambda_k)' = M\beta,$$

where $M = (x_{ij})_{i=1,\dots,k;j=0,\dots,n}$. We call a $k \times (n+1)$ matrix M as a *model matrix* of the log-linear model (1).

To judge the fitting of the log-linear model (1), we can consider various goodness-of-fit tests. In the goodness-of-fit tests, the model (1) is treated as the null model, whereas the saturated model is treated as the alternative model. Under the null model (1), the sufficient statistics for the (nuisance) parameter β is given by $M'\mathbf{y}$ from the factorization

$$\prod_{i=1}^{k} e^{-\lambda_i} \frac{\lambda_i^{y_i}}{y_i!} = \left(\prod_{i=1}^{k} \frac{1}{y_i!} \right) \exp \left(\beta' M' \mathbf{y} - \sum_{i=1}^{k} \lambda_i \right).$$

Therefore the conditional distribution of $\mathbf{y}$ for the given sufficient statistics is written as

$$f(\mathbf{y} \mid M'\mathbf{y} = M'\mathbf{y}^o) = C(M'\mathbf{y}^o)^{-1} \prod_{i=1}^{k} \frac{1}{y_i!}, \tag{2}$$

where $\mathbf{y}^o$ is the observation vector and $C(M'\mathbf{y}^o)$ is the normalizing constant determined from the sufficient statistics $M'\mathbf{y}^o$ as

$$C(M'\mathbf{y}^o) = \sum_{\mathbf{y} \in \mathcal{F}(M'\mathbf{y}^o)} \left(\prod_{i=1}^{k} \frac{1}{y_i!} \right) \tag{3}$$

and

$$\mathcal{F}(M'\mathbf{y}^o) = \{\mathbf{y} \in \mathbb{Z}^k_{\geq 0} \mid M'\mathbf{y} = M'\mathbf{y}^o\}. \tag{4}$$

In this paper, we consider goodness-of-fit tests based on the conditional distribution (2). There are several choices of the test statistics $T(\mathbf{y})$. Frequently used choices are the likelihood ratio statistics

$$T(\mathbf{y}) = 2\sum_{i=1}^{k} y_i \log \frac{y_i}{\hat{\lambda}_i} \tag{5}$$

or the Pearson χ^2 statistics

$$T(\mathbf{y}) = \sum_{i=1}^{k} \frac{(y_i - \hat{\lambda}_i)^2}{\hat{\lambda}_i},$$

where $\hat{\lambda}_i$ is the maximum likelihood estimate for λ_i under the null model (i.e., fitted value). A simple way of judging the significance for the observed value $T(\mathbf{y}^o)$ is the asymptotic p value based on the asymptotic distribution χ^2_{k-n-1} of the test statistics. However, the fitting of the asymptotic approximation may be sometimes poor. Therefore we consider conditional exact p values in this paper. Based on the conditional distribution (2), the exact conditional p value is written as

$$p = \sum_{\mathbf{y} \in \mathcal{F}(M'\mathbf{y}^o)} f(\mathbf{y} \mid M'\mathbf{y} = M'\mathbf{y}^o)\mathbf{1}(T(\mathbf{y}) \geq T(\mathbf{y}^o)), \tag{6}$$

where

$$\mathbf{1}(T(\mathbf{y}) \geq T(\mathbf{y}^o)) = \begin{cases} 1, & \text{if } T(\mathbf{y}) \geq T(\mathbf{y}^o), \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

is the test function of $T(\mathbf{y})$. Of course, if we can calculate the exact p value of (6) and (7), it is best. However, the cardinality of the set $\mathcal{F}(M'\mathbf{y}^o)$ becomes huge for moderate sizes of data and the calculation of the normalizing constant $C(M'\mathbf{y}^o)$ of (3) is usually computationally infeasible. Instead, we consider a Markov chain Monte Carlo method to evaluate the conditional p values. It should be noted that we need not calculate the normalizing constant (3) to evaluate the p values by the Markov chain Monte Carlo methods. This point is one of the important advantages of the Markov chain Monte Carlo methods.

To perform the Markov chain Monte Carlo procedure, we have to construct an irreducible Markov chain over the conditional sample space (4) with the stationary distribution (2). If such a chain is constructed, we can sample from the chain as $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(T)}$ after discarding some initial burn-in steps, and estimate the p values as

$$\hat{p} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}(T(\mathbf{y}^{(t)} \geq T(\mathbf{y}^o)).$$

Such a chain can be constructed easily by *Markov bases*. Once a Markov basis is obtained, we can construct a connected, aperiodic and reversible Markov chain over the conditional sample space (4), which can be modified so as to have the stationary distribution (2) by the Metropolis-Hastings procedure. See [11] or [10] for details.

The Markov basis is characterized algebraically as follows. Write the variables $x_1, \ldots, x_k$ and consider the polynomial ring $K[x_1, \ldots, x_k]$ over a field K. Consider the integer kernel of the transpose of the model matrix M, $\mathrm{Ker}_{\mathbb{Z}} M'$. For each $\mathbf{b} = (b_1, \ldots, b_k)' \in \mathrm{Ker}_{\mathbb{Z}} M'$, define a binomial in $K[x_1, \ldots, x_k]$ as

$$f_{\mathbf{b}} = \prod_{b_i > 0} x_i^{b_i} - \prod_{b_i < 0} x_i^{-b_i} .$$

Then the binomial ideal in $K[x_1, \ldots, x_k]$,

$$I_{M'} = \langle \{ f_{\mathbf{b}} \mid \mathbf{b} \in \mathrm{Ker}_{\mathbb{Z}} M' \} \rangle$$

is called a *toric ideal of the configuration* M'. Then for a generating set of $I_{M'}$, $\{f_{\mathbf{b}^{(1)}}, \ldots, f_{\mathbf{b}^{(s)}}\}$, the set of integer vectors $\{\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(s)}\}$ constitutes a Markov basis. See [10] for details.

3 Statistical models of the centrally symmetric configurations and Box-Behnken designs

As we have seen in Section 2, if we can obtain a generator of $I_{M'}$, a toric ideal of the configuration M', we can judge the fitting of the statistical model expressed by the model matrix M by the conditional p values estimated by the Markov chain Monte Carlo methods. For small sizes of problems, we can rely on various softwares such as 4ti2 ([1]) to compute generators of the toric ideals. However, for problems of large sizes, it is usually very difficult to compute Markov bases or Gröbner bases for given ideals. On the other hand, if we have theoretical results on the structure of the corresponding ideals, it is very easy to perform the Markov chain Monte Carlo procedure for such configurations. The *centrally symmetric configuration* is an example of such cases.

The centrally symmetric configuration is given in ([15]) as follows. Let $A \in \mathbb{Z}^{n \times s}$ be an integer matrix for which no column vector is a zero vector. Then the centrally symmetric configuration of A is the $(n+1) \times (2s+1)$ integer matrix

$$A^{\pm} = \left(\begin{array}{c|ccc|ccc} 0 & & & & & & \\ \vdots & & A & & & -A & \\ 0 & & & & & & \\ \hline 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{array} \right) . \tag{8}$$

It is known that the "toric ring" $K[A^{\pm}]$ of $A^{\pm}$ is normal and Gorenstein if there exists a squarefree initial ideal with respect to a reverse lexicographic order where

the smallest variable corresponds to the first column of $A^{\pm}$. See, e.g., [12, Lemma 1.1].

As natural statistical models and designs where (the transpose of) the model matrix is centrally symmetric configurations, we consider the simple first-order models as follows. Suppose $F \subset A_1 \times \cdots A_m \in \mathbb{Q}^m$ is a *symmetric design including the origin*, i.e., a design satisfying

$$(0,\ldots,0) \in F \ \text{ and } \ \mathbf{d} \in F \ \Rightarrow \ -\mathbf{d} \in F.$$

Write $D = (d_{ij}) \in \mathbb{Z}^{k\times m}$ its design matrix, where d_{ij} is the level of j-th factor in i-th run for $i = 1,\ldots,k,\ j = 1,\ldots,m$. Then we see that the transpose of the model matrix

$$M = \begin{pmatrix} 1 & \\ \vdots & D \\ 1 & \end{pmatrix} \tag{9}$$

is centrally symmetric. Corresponding log-linear model (1) is written as

$$\log \lambda_i = \beta_0 + \beta_1 d_{i1} + \cdots + \beta_m d_{im},\ i = 1,\ldots,k. \tag{10}$$

We call the model (10) as a *first-order model* in this paper. The interpretation of the first-order model (10) is as follows. Suppose there are adequate meanings both in the order of the factors and the interval of the factors for the design F. Then the model (10) means that the logarithm of the influence to the response variable is proportional to the difference of the levels for each factor. In other words, the parameter β_j represents the main effect of the j-th factor for $j = 1,\ldots,m$. The first order model (10) is a discrete version of the first-order model arising in the context of the *response surface methodology*. See Section 9 of [16], for example. A typical example of the symmetric designs is also arising in the context of the response surface methodology as *Box-Behnken designs*.

The Box-Behnken design is a family of three-level fractional factorial designs introduced by [9]. This design is constructed by combining two-level factorial designs with balanced (or partially balanced) incomplete block designs in a particular manner. To illustrate the concept of the Box-Behnken designs, consider the case of three factors (i.e., $m = 3$). A balanced incomplete block design with three factors and three blocks is given as follows.

	Factor		
Block	1	2	3
1	∘	∘	
2	∘		∘
3		∘	∘

The Box-Behnken design is constructed by replacing the two circles (∘) in each block by the two columns of the two-level 2^2 design and add a column of zeros

where a circle does not appear, and adding a run at the origin. In this example, the Box-Behnken design is constructed as follows.

Factor		
1	2	3
-1	-1	0
-1	1	0
1	-1	0
1	1	0
-1	0	-1
-1	0	1
1	0	-1
1	0	1
0	-1	-1
0	-1	1
0	1	-1
0	1	1
0	0	0

Similarly, by combining various incomplete block designs with two-level full (or fractional) factorial designs, various three-level fractional factorial designs are obtained. In this paper, we only consider the Box-Behnken designs constructed from the two-level 2^2 design and the balanced incomplete block designs with the block size 2, the number of factors (or treatments) m, the number of blocks $m(m-1)/2$ and the number of replicates for each factor $m-1$, with a single run at the origin. Note that it is common to consider the designs with several runs at the origins in this field. See [9] or Chapter 9 of [16] for details. However, we only consider the designs with single observations even in the origin. Therefore the Box-Behnken design considered in this paper has $2m(m-1)+1$ runs for m factor case.

For these Box-Behnken designs, we consider the first-order model (10) with the model matrix (9). Note that $n = m$ in our cases. Then we see that the transpose of the model matrix, M', has the centrally symmetric structure (8) with $s = m(m-1)$. For example, the transpose of the model matrix of the first-order model for the three factors case is given by

$$\left(\begin{array}{rrrrrrrrrrrrr} 0 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right),$$

which is the centrally symmetric configuration of

$$\left(\begin{array}{rrrrrr} -1 & -1 & -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}\right).$$

Following the arguments of Section 2, we can judge the fitting of the first-order model for the Box-Behnken designs by the Markov chain Monte Carlo methods, if we obtain the generators of the toric ideal of the configuration of this type.

4 Gröbner bases of centrally symmetric configurations of root system D_n

Now we show the structure of the Gröbner bases of the centrally symmetric configurations for the Box-Behnken designs. Because the Gröbner basis is a generator of the ideals, we can use the Gröbner basis as a Markov basis. As an important fact, the transpose of the model matrix for the first-order models for the Box-Behnken designs is characterized as the configuration of the *root system* D_n.

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ stand for the canonical unit coordinate vectors of $\mathbb{R}^n$ and $\mathbf{D}_n^{\pm} \subset \mathbb{R}^n$ the finite set which consists of the origin $\mathbf{0}$ of $\mathbb{R}^n$ together with

$$\mathbf{e}_i + \mathbf{e}_j,\ \mathbf{e}_i - \mathbf{e}_j,\ -\mathbf{e}_i + \mathbf{e}_j,\ -\mathbf{e}_i - \mathbf{e}_j, \quad 1 \leq i < j \leq n.$$

Let $K[\mathbf{t}, \mathbf{t}^{-1}, s] = K[t_1, \ldots, t_n, t_1^{-1}, \ldots, t_n^{-1}, s]$ denote the Laurent polynomial ring in $2n+1$ variables over a field K. The *toric ring* of $\mathbf{D}_n^{\pm}$ is the subring $K[\mathbf{D}_n^{\pm}]$ of $K[\mathbf{t}, \mathbf{t}^{-1}, s]$ which is generated by s together with $t_i t_j s, t_i t_j^{-1} s, t_i^{-1} t_j s, t_i^{-1} t_j^{-1} s$, where $1 \leq i < j \leq n$. Let $K[\mathbf{x}, z]$ be the polynomial ring over K in the variables z together with x_{ij}^{pq}, where $1 \leq i < j \leq n$ and $p, q \in \{+, -\}$. We then define the surjective ring homomorphism $\pi : K[\mathbf{x}, z] \to K[\mathbf{D}_n^{\pm}]$ by setting $\pi(z) = s$ and $\pi(x_{ij}^{pq}) = t_i^{p'} t_j^{q'} s$, where $+' = 1$ and $-' = -1$. For example $\pi(x_{25}^{-+}) = t_2^{-1} t_5 s$. The *toric ideal* of $\mathbf{D}_n^{\pm}$ is the kernel $I_{\mathbf{D}_n^{\pm}}$ of π.

Fix an ordering $<$ of the variables of $K[\mathbf{x}]$ with the property that $x_{ij}^{pq} < x_{k\ell}^{rs}$ if either (i) $i < k$ or (ii) $i = k$ and $j > \ell$. Let $<_{\text{lex}}$ denote the lexicographic order on $K[\mathbf{x}]$ induced by the ordering $<$. We introduce the monomial order $\prec$ on $K[\mathbf{x}, z]$ defined as follows: One has $\prod_{\xi=1}^{a} x_{i_\xi j_\xi}^{p_\xi q_\xi} z^\alpha \prec \prod_{\nu=1}^{b} x_{k_\nu \ell_\nu}^{r_\nu s_\nu} z^\beta$, where $\alpha, \beta \in \mathbb{Z}_{\geq 0}$, if

- $a + \alpha < b + \beta$, or
- $a + \alpha = b + \beta$ and $\alpha > \beta$, or
- $a = b, \alpha = \beta$ and $\prod_{\xi=1}^{a} x_{i_\xi j_\xi}^{p_\xi q_\xi} <_{\text{lex}} \prod_{\nu=1}^{b} x_{k_\nu \ell_\nu}^{r_\nu s_\nu}$.

Let $\mathcal{G}$ denote the set of binomials

(i) $x_{ij}^{pq} x_{k\ell}^{rs} - x_{ik}^{pr} x_{j\ell}^{qs}, \quad 1 \leq i < j < k < \ell \leq n;$

(ii) $x_{i\ell}^{ps} x_{jk}^{qr} - x_{ik}^{pr} x_{j\ell}^{qs}, \quad 1 \leq i < j < k < \ell \leq n;$

(iii) $x_{ij}^{+p} x_{ik}^{-q} - x_{jk}^{pq} z, \quad |\{i, j, k\}| = 3;$[1]

[1] For $j < i$, the notation x_{ij}^{pq} is identified with the variable x_{ji}^{qp}.

(iv) $x_{ij}^{++}x_{ij}^{--} - z^2, \ \ 1 \leq i < j \leq n;$

(v) $x_{ij}^{+-}x_{ij}^{-+} - z^2, \ \ 1 \leq i < j \leq n;$

(vi) $x_{ij}^{p+}x_{ij}^{p-} - x_{1i}^{+p}x_{1i}^{-p}, \ \ 1 < i \neq j \leq n;$

(vii) $x_{1j}^{p+}x_{1j}^{p-} - x_{1n}^{p+}x_{1n}^{p-}, \ \ 1 < j < n;$

(viii) $x_{1i}^{+p}x_{1i}^{-p}x_{jk}^{qr} - x_{ij}^{pq}x_{ik}^{pr}z, \ \ 1 < i, j, k \leq n, |\{i, j, k\}| = 3$

belonging to $I_{\mathbf{D}_n^\pm}$.

Theorem 4.1. *The set $\mathcal{G}$ of binomials is a Gröbner basis of $I_{\mathbf{D}_n^\pm}$ with respect to $\prec$.*

Proof. In general, if $f = u - v$ is a binomial, then u is called the *first* monomial of f and v is called the *second* monomial of f. The initial monomial of each of the binomials (i) – (viii) with respect to $\prec$ is its first monomial. Let $\mathrm{in}_\prec(\mathcal{G})$ denote the set of initial monomials of binomials belonging to $\mathcal{G}$. It follows from [14, (0.1)] that, in order to show that $\mathcal{G}$ is a Gröbner basis of $I_{\mathbf{D}_n^\pm}$ with respect to $\prec$, what we must prove is the following: (♣) If u and v are monomials belonging to $K[\mathbf{x}, z]$ with $u \neq v$ such that $u \notin \mathrm{in}_\prec(\mathcal{G})$ and $v \notin \mathrm{in}_\prec(\mathcal{G})$, then $\pi(u) \neq \pi(v)$.

Let u and v be monomials belonging to $K[\mathbf{x}, z]$. Write

$$u = x_{i_1j_1}^{p_1q_1} \cdots x_{i_aj_a}^{p_aq_a} z^\alpha, \qquad v = x_{k_1\ell_1}^{r_1s_1} \cdots x_{k_b\ell_b}^{r_bs_b} z^\beta$$

with

$$i_1 \leq \cdots \leq i_a, \qquad k_1 \leq \cdots \leq k_b.$$

Let $\pi(u) = \pi(v)$. Then $a + \alpha = b + \beta$. Suppose that $u \notin \mathrm{in}_\prec(\mathcal{G})$ and $v \notin \mathrm{in}_\prec(\mathcal{G})$. Furthermore, suppose that u and v are relatively prime. Especially either $\alpha = 0$ or $\beta = 0$. Let, say, $\alpha = 0$. In other words,

$$u = x_{i_1j_1}^{p_1q_1} \cdots x_{i_aj_a}^{p_aq_a}, \qquad v = x_{k_1\ell_1}^{r_1s_1} \cdots x_{k_b\ell_b}^{r_bs_b} z^\beta,$$

where $\beta = a - b$.

Let $i_{a'} < i_{a''}$, where $1 \leq a' < a'' \leq a$. Then, by using (i), one has $i_{a''} \leq j_{a'}$ and, by using (ii), one has $j_{a'} \leq j_{a''}$. Hence $i_{a'} < i_{a''} \leq j_{a'} \leq j_{a''}$. It then follows that

$$i_1 \leq \cdots \leq i_a \leq j_1 \leq \cdots \leq j_a, \qquad k_1 \leq \cdots \leq k_b \leq \ell_1 \leq \cdots \leq \ell_b. \tag{11}$$

We claim that none of the followings arises:

(♯) $i_{a'} = i_{a''}$ and $p_{a'} \neq p_{a''}$;

(♭) $i_a = j_1$ and $p_a \neq q_1$;

(♮) $j_{a'} = j_{a''}$ and $q_{a'} \neq q_{a''}$.

(Case ($\sharp$)) Let $i_{a'} = i_{a''}$ and $p_{a'} \neq p_{a''}$, where $1 \leq a' < a'' \leq a$. Then, by using (iii), one has $j_{a'} = j_{a''}$. Then, by using (iv) and (v), one has $q_{a'} = q_{a''}$. Moreover, by using (vi), one has $i_{a'} = i_{a''} = 1$. Thus $x_{i_{a'}j_{a'}}^{p_{a'}q_{a'}} x_{i_{a''}j_{a''}}^{p_{a''}q_{a''}} = x_{1j_{a'}}^{+q_{a'}} x_{1j_{a'}}^{-q_{a'}}$ divides u. Then, by using (viii), each variable $x_{i_{a^*}j_{a^*}}^{p_{a^*}q_{a^*}}$ ($\neq x_{1j_{a'}}^{+q_{a'}}, \neq x_{1j_{a'}}^{-q_{a'}}$) which divides u satisfies either $i_{a^*} = 1$ or $|\{j_{a'}, i_{a^*}, j_{a^*}\}| = 2$ with $i_{a^*} > 1$. Then, by using (iii), if $i_{a^*} = 1$, then $j_{a'} = j_{a^*}$. Thus $q_{a'} \neq q_{a^*}$. However, by using (iv) and (v), a contradiction arises. Hence either $j_{a'} = i_{a^*} > 1$ or $j_{a'} = j_{a^*} > 1$. In other words, u is divided by either $x_{1j_{a'}}^{+q_{a'}} x_{1j_{a'}}^{-q_{a'}} x_{j_{a'}j_{a^*}}^{p_{a^*}q_{a^*}}$ or $x_{1j_{a'}}^{+q_{a'}} x_{1j_{a'}}^{-q_{a'}} x_{i_{a^*}j_{a'}}^{p_{a^*}q_{a^*}}$. Then, by using (iii), if $j_{a'} = i_{a^*}$, then $q_{a'} = p_{a^*}$. Again, by using (iii), if $j_{a'} = j_{a^*}$, then $q_{a'} = q_{a^*}$. Hence either $x_{i_{a^*}j_{a^*}}^{p_{a^*}q_{a^*}} = x_{j_{a'}j_{a^*}}^{q_{a'}q_{a^*}}$ or $x_{i_{a^*}j_{a^*}}^{p_{a^*}q_{a^*}} = x_{i_{a^*}j_{a'}}^{p_{a^*}q_{a'}}$. As a result, either $t_{j_{a'}}^{a}$ or $t_{j_{a'}}^{-a}$ divides $\pi(u)$. Let, say, $t_{j_{a'}}^{a}$ divides $\pi(u)$. Since $\pi(u) = \pi(v)$, it follows that $\beta = a - b = 0$ and that $t_{j_{a'}}^{a}$ divides $\pi(v)$.

Let, say, either t_j or t_j^{-1}, where $j \neq j_{a'}$, divide $\pi(u) = \pi(v)$. Then either $x_{j_{a'}j}^{++}$ or $x_{j_{a'}j}^{+-}$ must divide both u and v, which contradicts the fact that u and v are relatively prime. Hence $\pi(u) = \pi(v) = t_{j_{a'}}^{a} s^a$. Thus a quadratic monomial $x_{j_{a'}j_0}^{++} x_{j_{a'}j_0}^{+-}$, where $j_0 \neq j_{a'}$, divides u. Then, since $j_{a'} > 1$, by using (vi), one has $j_0 = 1$. Hence $x_{1j_{a'}}^{++} x_{1j_{a'}}^{-+}$ divides both u and v, a contradiction.

(Case ($\flat$)) Let $i_a = j_1$ and $p_a \neq q_1$. Then, by using (iii), one has $u \in \mathrm{in}_{\prec}(\mathcal{G})$.

(Case ($\natural$)) Let $j_{a'} = j_{a''}$ and $q_{a'} \neq q_{a''}$, where $1 \leq a' < a'' \leq a$. Then, by using (iii), one has $i_{a'} = i_{a''}$. Furthermore, by using (iv) and (v), one has $p_{a'} = p_{a''}$. If $i_{a'} > 1$, then u is divided by $x_{i_{a'}j_{a'}}^{p_{a'}q_{a'}} x_{i_{a''}j_{a''}}^{p_{a''}q_{a''}} = x_{i_{a'}j_{a'}}^{p_{a'}q_{a'}} x_{i_{a'}j_{a'}}^{p_{a'}-q_{a'}}$. Thus, by using (vi), one has $u \in \mathrm{in}_{\prec}(\mathcal{G})$. Hence $i_{a'} = i_{a''} = 1$. Thus, by using (vii), one has $j_{a'} = j_{a''} = n$.

Let $i_{a^*} > 1$ for some $1 \leq a^* \leq n$. Then $a' < a^*$ and $1 = i_{a'} < i_{a^*} < n = j_{a'} \leq j_{a^*}$. Hence $j_{a'} = j_{a^*} = n$. However, since $i_{a^*} > 1$, it follows that $q_{a'} = q_{a^*}$. Thus $x_{i_{a''}j_{a''}}^{p_{a''}q_{a''}} x_{i_{a^*}j_{a^*}}^{p_{a^*}q_{a^*}} = x_{1n}^{p_{a'}-q_{a'}} x_{i_{a^*}n}^{p_{a^*}q_{a'}}$. Since $i_{a^*} < n$, by using (iii), one has $u \in \mathrm{in}_{\prec}(\mathcal{G})$. As a result, $i_{a^*} = 1$ for all $1 \leq a^* \leq a$.

Now, since ($\sharp$) cannot occur, one has $p_1 = p_{a^*}$ for all $1 \leq a^* \leq a$. Thus $\pi(u)$ is divided by either t_1^a or t_1^{-a}. Let, say, t_1^a divide $\pi(u)$. Since $\pi(u) = \pi(v)$, it follows that $\beta = a - b = 0$ and t_1^a divides $\pi(v)$. Let, say, either t_j or t_j^{-1}, where $j \neq 1$, divide $\pi(u) = \pi(v)$. Then either x_{1j}^{++} or x_{1j}^{+-} must divide both u and v, which contradicts the fact that u and v are relatively prime. Hence $\pi(u) = \pi(v) = t_1^a s^a$. Thus a quadratic monomial $x_{1j_0}^{++} x_{1j_0}^{+-}$, where $j_0 \neq 1$, divides u. Then, by using (vii), one has $j_0 = n$. Thus $x_{1n}^{++} x_{1n}^{+-}$ divides both u and v, a contradiction.

Finally, since none of ($\sharp$), ($\flat$) and ($\natural$) arises, it follows that no cancellation occurs in the expression of the Laurent monomial

$$\pi(u) = t_{i_1}^{p_1'} \cdots t_{i_a}^{p_a'} t_{j_1}^{q_1'} \cdots t_{j_a}^{q_a'}.$$

Since $\pi(u) = \pi(v)$, one has $\beta = a - b = 0$. Furthermore, since no cancellation occurs in the expression of the Laurent monomial

$$\pi(v) = t_{k_1}^{r_1'} \cdots t_{k_a}^{r_a'} t_{\ell_1}^{s_1'} \cdots t_{\ell_a}^{s_a'},$$

it follows from (11) that

$$i_\xi = k_\xi,\, j_\xi = \ell_\xi,\, p_\xi = r_\xi,\, q_\xi = s_\xi, \quad 1 \le \xi \le a.$$

Recall that u and v are relatively prime. Hence $a = 0$ and $u = v = 1$. Consequently the required condition (♣) is satisfied. $\square$

Example 4.2. Consider the case of $n = 3$. From Theorem 4.1,

$$\begin{array}{rl}
\text{(iii)} & x_{12}^{++}x_{13}^{-+} - x_{23}^{++}z,\;\; x_{12}^{++}x_{13}^{--} - x_{23}^{+-}z,\;\; x_{12}^{+-}x_{13}^{-+} - x_{23}^{-+}z,\;\; x_{12}^{+-}x_{13}^{--} - x_{23}^{--}z,\\
& x_{12}^{-+}x_{13}^{++} - x_{23}^{++}z,\;\; x_{12}^{-+}x_{13}^{+-} - x_{23}^{+-}z,\;\; x_{12}^{--}x_{13}^{++} - x_{23}^{-+}z,\;\; x_{12}^{--}x_{13}^{+-} - x_{23}^{--}z,\\
& x_{12}^{++}x_{23}^{-+} - x_{13}^{++}z,\;\; x_{12}^{++}x_{23}^{--} - x_{13}^{+-}z,\;\; x_{12}^{-+}x_{23}^{-+} - x_{13}^{-+}z,\;\; x_{12}^{-+}x_{23}^{--} - x_{13}^{--}z,\\
& x_{12}^{+-}x_{23}^{++} - x_{13}^{++}z,\;\; x_{12}^{+-}x_{23}^{+-} - x_{13}^{+-}z,\;\; x_{12}^{--}x_{23}^{++} - x_{13}^{-+}z,\;\; x_{12}^{--}x_{23}^{+-} - x_{13}^{--}z,\\
& x_{13}^{++}x_{23}^{+-} - x_{12}^{++}z,\;\; x_{13}^{++}x_{23}^{--} - x_{12}^{+-}z,\;\; x_{13}^{-+}x_{23}^{+-} - x_{12}^{-+}z,\;\; x_{13}^{-+}x_{23}^{--} - x_{12}^{--}z,\\
& x_{13}^{+-}x_{23}^{++} - x_{12}^{++}z,\;\; x_{13}^{+-}x_{23}^{-+} - x_{12}^{+-}z,\;\; x_{13}^{--}x_{23}^{++} - x_{12}^{-+}z,\;\; x_{13}^{--}x_{23}^{-+} - x_{12}^{--}z,\\
\text{(iv)} & x_{12}^{++}x_{12}^{--} - z^2,\;\; x_{13}^{++}x_{13}^{--} - z^2,\;\; x_{23}^{++}x_{23}^{--} - z^2,\\
\text{(v)} & x_{12}^{+-}x_{12}^{-+} - z^2,\;\; x_{13}^{+-}x_{23}^{-+} - z^2,\;\; x_{23}^{+-}x_{13}^{-+} - z^2,\\
\text{(vi)} & x_{23}^{++}x_{23}^{+-} - x_{12}^{++}x_{12}^{-+},\;\; x_{23}^{-+}x_{23}^{--} - x_{12}^{+-}x_{12}^{--},\\
& x_{23}^{++}x_{23}^{-+} - x_{13}^{++}x_{13}^{-+},\;\; x_{23}^{+-}x_{23}^{--} - x_{13}^{+-}x_{13}^{--},\\
\text{(vii)} & x_{12}^{++}x_{12}^{+-} - x_{13}^{++}x_{13}^{+-},\;\; x_{12}^{-+}x_{12}^{--} - x_{13}^{-+}x_{13}^{--}
\end{array}$$

is the Gröbner basis $\mathcal{G}$ of $I_{\mathbf{D}_n^{\pm}}$ with respect to $\prec$.

We have the following since the binomial $x_{1i}^{+p}x_{1i}^{-p}x_{jk}^{qr} - x_{ij}^{pq}x_{ik}^{pr}z$ (viii) where $1 < i, j, k \le n$ and $|\{i, j, k\}| = 3$ satisfies

$$\begin{aligned}
x_{1i}^{+p}x_{1i}^{-p}x_{jk}^{+r} - x_{ij}^{p+}x_{ik}^{pr}z &= x_{jk}^{+r}(x_{1i}^{+p}x_{1i}^{-p} - x_{ij}^{p+}x_{ij}^{p-}) + x_{ij}^{p+}(x_{ij}^{p-}x_{jk}^{+r} - x_{ik}^{pr}z),\\
x_{1i}^{+p}x_{1i}^{-p}x_{jk}^{-r} - x_{ij}^{p-}x_{ik}^{pr}z &= x_{jk}^{-r}(x_{1i}^{+p}x_{1i}^{-p} - x_{ij}^{p+}x_{ij}^{p-}) + x_{ij}^{p-}(x_{ij}^{p+}x_{jk}^{-r} - x_{ik}^{pr}z).
\end{aligned}$$

Corollary 4.3. *The toric ideal $I_{\mathbf{D}_n^{\pm}}$ is generated by binomials* (i) – (vii) *in $\mathcal{G}$ in Theorem* 4.1. *In particular, $I_{\mathbf{D}_n^{\pm}}$ is generated by quadratic binomials.*

5 Numerical example

In this section, we perform our Markov chain procedure to an imaginary data set. Our data set is constructed from actual experimental data as follows. In [13], the Box-Behnken design is used to apply the response surface method. The purpose of this experiment is to determine the optimal processing condition of a pulsed UV-light system to inactivate the fungal spores of Aspergillus niger in corn meal. The three factors are A: Treatment time (20, 60, 100 second), B: Distance from the UV

strobe (3, 8, 13 cm), and C: Voltage input (2000, 2900, 3800 V). The response is the reduction of the Aspergillus niger in the $\log_{10}$ scale. In [13], the first-order and the second-order polynomial models for the response are considered. See [13] for detail description of the data analysis. Because our method is for discrete data, we use the rounded values of (10 times of) the responses in this experimental data and treat them as realizations of discrete variables. Then we have an imaginary data set in Table 1. Because the responses in the original data in [13] are continuous values,

Table 1: Box-Behnken design matrix for the three factors and the response (The responses are rounded values of Table 1 of [13]. Fitted values are calculated under the model (10))

Time (s)	Distance (cm)	Voltage (V)	Response	Fitted values
20	3	2900	4	4.04
20	13	2900	3	3.59
100	3	2900	33	31.58
100	13	2900	30	28.02
20	8	2000	2	2.12
20	8	3800	5	6.84
100	8	2000	14	16.57
100	8	3800	50	53.42
60	3	2000	7	6.29
60	3	3800	21	20.29
60	13	2000	5	5.58
60	13	3800	20	18.01
60	8	2900	13	10.64

we cannot emphasize our computational results from the applied statistical view. The purpose of this numerical experiment is only to check that our method works for some discrete data. However, it can also be natural to consider the fitting of the log-linear model (1) to the response because the original response is reported in $\log_{10}$ scale.

For the response data in Table 1, we consider the fitting of the first-order model (10) based on the likelihood ratio statistics (5). The fitted values under the null model is calculated in the last column of Table 1. The likelihood ratio is 2.36 with 9 degree of freedom. Therefore the asymptotic p value is 0.98 from the asymptotic χ_9^2 distribution. To evaluate the fitting of the first-order model (10), we perform the Markov chain Monte Carlo method. We use the Gröbner basis given in Example 4.2 as a Markov basis. After 50000 burn-in steps from the observed data as the initial state, we derive 100000 samples by Metropolis-Hasting algorithm. Among these samples, 98834 samples have the larger likelihood ratio values than the observed 2.36. Therefore the conditional p value is estimated as 0.99, which suggests the

good fitting of the first-order model (10). Figure 1 is the histogram of the sample likelihood ratio statistics with the asymptotic χ^2_9 distribution.

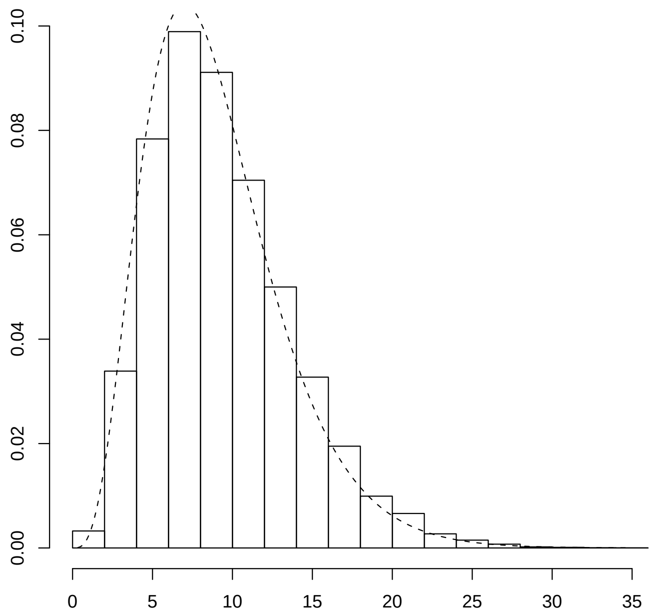

Figure 1: Asymptotic and Monte Carlo estimated distribution of the likelihood ratio statistics

6 Discussion

In this paper, we present a new method for analyzing non-regular fractional factorial designs. The motivation of this paper is a new finding on the structure of the Gröbner bases of the centrally symmetric configurations of root system D_n. As we have seen in the paper, we can relate the theoretical results in the algebraic field to the statistical problems for the Box-Behnken designs. Our model is simple and fundamental. In fact, we usually consider more complicated models such as second-order model for the analysis of the Box-Behnken designs. However, the structure of the Markov bases or the Gröbner bases for the second-order model is very complicated. Though the Markov chain Monte Carlo methods can be considered for general non-regular designs, the structure of the Markov bases is only revealed for simple models such as the hierarchical models for the regular designs at present. Therefore we think our contribution on the new results of the non-regular designs is

important. Besides, compared to the continuous data analysis on the assumption of the normality, there are very few experiments are reported treating the discrete data arising in the fractional factorial designs. We think our Markov chain Monte Carlo procedure is very simple and can be used easily, and can be one of the powerful choices in the analysis of the discrete data.

References

[1] 4ti2 team. 4ti2 – A software package for algebraic, geometric and combinatorial problems on linear spaces. Available at `www.4ti2.de`.

[2] S. Aoki, H. Hara and A. Takemura (2012). *Markov bases in algebraic statistics*. Springer Series in Statistics.

[3] S. Aoki, T. Hibi and H. Ohsugi (2013). Markov chain Monte Carlo methods for the regular two-level fractional factorial designs and cut ideals. *J. Statist. Plann. Infer.*, **143**, 1791–1806.

[4] S. Aoki, T. Hibi, H. Ohsugi and A. Takemura (2008). Gröbner bases of nested configurations. *J. Algebra*, **320**, 2583–2593.

[5] S. Aoki, T. Hibi, H. Ohsugi and A. Takemura (2010). Markov basis and Gröbner basis of Segre-Veronese configuration for testing independence in group-wise selections. *Ann. Inst. Statist. Math.*, **62**(2), 299-321.

[6] S. Aoki and A. Takemura (2009). Markov basis for design of experiments with three-level factors. in *Algebraic and Geometric Methods in Statistics* (dedicated to Professor Giovanni Pistone on the occasion of his sixty-fifth birthday), edited by P. Gibilisco, E. Riccomagno, M. P. Rogantin and H. P. Wynn, Cambridge University Press, 225–238.

[7] S. Aoki and A. Takemura (2010). Markov chain Monte Carlo tests for designed experiments. *Journal of Statistical Planning and Inference*, **140**, 817–830.

[8] F. Ardila, M. Beck, S. Hoşten, J. Pfeifle, and K. Seashore (2011). Root Polytopes and Growth Series of Root Lattices, *SIAM J. Discrete Math.*, **25** (1), 360–378.

[9] G. E. P. Box and D. W. Behnken (1960). Some new three level designs for the study of quantitative variables. *Technometrics*, **2**, 455–475.

[10] P. Diaconis and B. Sturmfels (1998). Algebraic algorithms for sampling from conditional distributions. *Annals of Statistics*, **26**, 363–397.

[11] W. K. Hastings (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

[12] T. Hibi, K. Matsuda, H. Ohsugi, and K. Shibata (2014). Centrally symmetric configurations of order polytopes, preprint. (`arXiv:1409.4386 [math.AC]`.)

[13] S. Jun, J. Irudayaraj, A. Demirci and D. Geiser (2003). Pulsed UV-light treatment of corn meal for inactivation of Aspergillus niger spores. *International Journal of Food Science and Technology*, **38**, 883–888.

[14] H. Ohsugi and T. Hibi (2002). Quadratic initial ideals of root systems, *Proc. Amer. Math. Soc.*, **130**, 1913–1922.

[15] H. Ohsugi and T. Hibi (2014). Centrally symmetric configurations of integer matrices. *Nagoya Math. J.*, in press (doi:10.1215/00277630-2857555).

[16] C. F. J. Wu and M. Hamada (2000). *Experiments. Planning, Analysis, and Parameter Design Optimization*. John Wiley & Sons.